

# 動的情報が複数の視覚情景刺激提示下の視聴覚統合に及ぼす影響

小西 慶治  
横澤 一彦

東京大学大学院人文社会系研究科

東京大学大学院人文社会系研究科

視覚刺激と聴覚刺激を一对一で用いた腹話術効果の研究は、視聴覚統合における時空間的近接性の影響を調べるに過ぎないが、多対一では意味的整合性まで考慮する必要がある (Kanaya & Yokosawa, 2011, *Psychonomic Bulletin & Review*). 本研究では、視覚情景刺激2つと聴覚情景刺激1つを用いて、視聴覚刺激の意味的整合性と視覚刺激の動き情報とが視聴覚統合へ及ぼす影響を調べた。動画と静止画を左右に並べ、聴覚刺激との意味的整合性を操作した。聴覚刺激は左右いずれかから提示し、参加者はどちらから提示されたかを回答した。すると、動画が意味的に整合する場合、不整合の場合よりも音源定位が特に不正確になり、静止画が意味的に整合していてもそれほど不正確にならなかった。これらから、複雑な環境における視聴覚の対応づけでは、動き情報を持つ視覚刺激との意味的整合性が優先され、動き情報を持たない視覚刺激との意味的整合性の影響は小さいことが分かった。

Keywords: cross-modal integration, the spatial ventriloquist effect, semantic congruency

## 問題・目的

聴覚刺激と視覚刺激を空間的に離れた位置に提示すると、聴覚刺激の音源位置が視覚刺激の位置に誤定位される現象は、腹話術効果 (the ventriloquist effect) と呼ばれる (Bertelson & Ascherleben, 1998)。これは、感覚入力間の矛盾を解消し、一貫した外界の表象を形成する統合処理の典型例である。聴覚刺激と視覚刺激を1つずつ提示し、それらの矛盾を操作した実験によると、時間的近接性のような物理的要因こそが統合の成否を規定すると考えられる (Vroomen & de Gelder, 2004)。一方で、Kanaya & Yokosawa (2011) は、顔刺激を2つ、声刺激1つを用いて音源定位課題を行い、視聴覚情報の意味的な対応関係が統合の程度を調整することを報告した。これは、視聴覚情報の対応づけ候補が複数存在する環境下では、認知的要因に基づいて、蓋然性の高い対応を選択することを示唆している。

しかしながら、Kanaya & Yokosawa (2011) では、顔刺激の一方の口元を隠すことで対応づけに関わる情報をなく奪っていたため、実質、口元の見える顔刺激と声刺激との対応関係のみを考えていた可能性がある。また、視聴覚統合における認知要因の影響はヒトの顔と声に特異的であるとの主張も存在する (Vatakis & Spence, 2008)。そこで本研究では、情景を表す視覚刺激を2枚と曖昧な聴覚刺激1つを提示し、視覚刺激の動き情報の有無と視聴覚情報の意味的整合性を操作した上での音源定位課題を行うことで、複数の対応関係の候補を残した複雑な環境下で、認知的要因が情景の視聴覚統合に影響するかを検討した。動的情報を有していても意味的に整合しなければ、あるいは、意味的に整合していても動的情報が伴わなければ音源の誤定位は生じないと考えられる。逆に、意味的に整合し、かつ、動的情報を伴う視覚刺激のみが聴覚刺激と強く統合されると予想される。

## 方法

**実験参加者** 実験には19名が参加した。

**刺激** 視覚刺激は炎、滝、雨、雪の4種類の情景を写した1秒間の動画であった。この動画の1フレーム目を用い、静止画とした。聴覚刺激はラジオ音を模した曖昧な雑音1種類であった。別の7名を対象に7件法で視覚刺激と聴覚刺激の整合度を評定させたところ、音との整合度は、雨 (Med = 3) や雪 (3) と比べて、炎 (6) や滝 (6) において高いことがわかったので、前者二つを整合性低刺激、後者二つを整合性高刺激と設定した。

**装置** 参加者の正面にスクリーンを設置し、プロジェクターを通じて視覚刺激を投影した。スクリーン前方下部に設置したスピーカーから聴覚刺激を提示した。

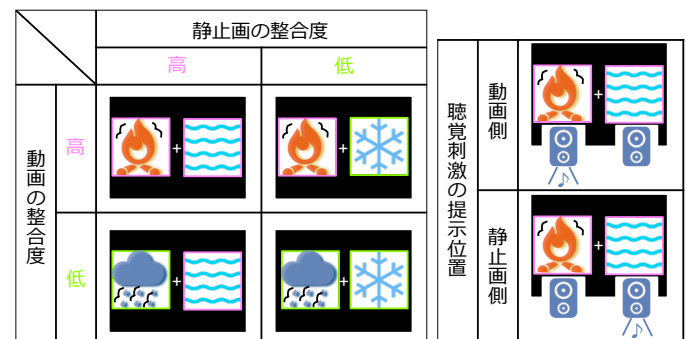


図1. 視覚刺激の提示例と聴覚刺激の提示例。

**実験手続き** 中央に注視点を提示したのち、その両側に視覚刺激を対提示すると同時に左右いずれか一方から聴覚刺激を提示した。課題は、聴覚刺激がどちらから提示されたかを答えることであった。参加者には、刺激の提示中、注視点を見つめるよう教示し、音源定位は視覚情報に頼らず、聴覚のみに基づくよう求めた。

**条件** 視覚刺激対のうち、一方は動画で、もう片方は静止画であった。4種類の情景から重複なく選ばれた。それぞれの視覚刺激について、聴覚刺激との意味的整合性が高い場合と低い場合の2通りが存在したので、視覚刺激対の提示条件は4通りだった(図1左)。また、聴覚刺激は視覚刺激対の動画側または静止画側から提示された(図1右)ので、視聴覚刺激の組み合わせは8条件からなっていた。各32試行ずつ、計256試行の音源判断が行われた。

## 結果

音源定位課題の正答率を表1に示す。聴覚刺激の提示位置(動画側/静止画側)、動画の意味的整合性(高/低)、静止画の意味的整合性(高/低)が、音源定位課題の正答率に及ぼす影響を検討するため、3要因反復測定分散分析を行なったところ、聴覚刺激の主効果 [ $F(1, 18) = 21.6, p < .001$ ], 動画の意味的整合性 [ $F(1, 18) = 22.7, p < .001$ ], 静止画の意味的整合性 [ $F(1, 18) = 10.0, p = .006$ ] の主効果がそれぞれ有意であった。すなわち、音源定位は静止画側よりも動画側で正確になり、動画との意味的整合度が高い場合に音源定位が不正確となる一方、静止画との意味的整合度が高い場合には音源定位が正確になっていた。

また、聴覚刺激の提示位置と動画の意味的整合性の二次の交互作用が有意であった [ $F(1, 18) = 40.0, p < .001$ ]。下位検定を行なったところ、提示位置の単純主効果は動画の整合度によらず認められた [高,  $F(1, 36) = 40.9, p < .001$ ; 低,  $F(1, 36) = 5.99, p = .02$ ]。さらに、動画の整合性の単純主効果は静止画側で認められたが、動画側では認められなかった [静止画側,  $F(1, 36) = 62.1, p < .001$ ; 動画側,  $F(1, 36) = 2.10, p = .16$ ]。聴覚刺激の提示位置と静止画の意味的整合性の交互作用をはじめ、その他の二次、および、三次の交互作用は有意ではなかった [ $F_s < 4.03, p_s > .060$ ]。これらから、動画との意味的整合度が高いと、実際には音源が静止画側に提示されていたにもかかわらず、動画側から提示されたと感じやすいが、静止画との意味的整合度は、音源の誤定位を生じないことがわかった。

表1. 音源定位の課題の正答率 (%)

		静止画の意味的整合性		
		高	低	
動画側から聴覚刺激が提示	動画の	高	97.2 ± 0.7	92.3 ± 1.5
	意味的整合性	低	90.6 ± 2.4	92.1 ± 2.2
		静止画側から聴覚刺激が提示		
		高	低	
静止画側から聴覚刺激が提示	動画の	高	61.8 ± 6.9	57.3 ± 7.0
	意味的整合性	低	81.0 ± 3.8	74.8 ± 5.8

Note: mean ± s.e.

## 考察

情景に関わる視覚刺激と聴覚刺激とを用いて、それらの意味的整合性および、視覚刺激が動きの情報を有しているかどうかといった認知的要因が視聴覚統合を調整するかどうかを調べた。音源定位の判断成績から、音源が静止画側に提示された場合に、動画側に提示された場合と比べて正答率が低く、動的情報を有している視覚刺激と同じ側から音が提示されたと感じられやすいことがわかる。これは、音が鳴る際には何らかの動きが伴う蓋然性が高いという事前の知識に基づくものである (Schutz & Kubovy, 2009)。動的な刺激と同側に提示された音源定位の成績が精確になるというのは研究とも一致する (Kanaya & Yokosawa, 2011)。

意味的な整合性の高い動画が提示された条件では、整合性の低い動画が提示された条件と比べて、静止画側に提示された聴覚刺激の定位成績の低下が著しかった。これは意味的に整合し、動きを伴う視覚刺激に対して聴覚刺激が強く統合されたため、実際には反対側に提示されたにもかかわらず、動画側から提示されたように感じられたと考えられる。

一方で、静止画が提示された場合は、このような音源定位成績の低下は観察されなかった。したがって、動的情報を伴う刺激の存在下では、たとえ意味的な整合性が高かったとしても動的情報の欠如を理由に適切な対応づけと判断されない可能性が考えられる。むしろ、静止画の意味的整合性が高い場合には、静止画側だけでなく、動画側においても正答率の向上が認められることから裏付けられる。この静止画の意味的整合性の効果は腹話術効果ではなく、単に意味的に不整合な視覚刺激によって聴覚刺激の処理が妨害されたことが原因だと考えられる。

以上のことから、意味的な整合性の視聴覚統合への寄与は限定的であり、視覚情報が動きを伴う場合に限ることが分かった。これは、認知的要因が統合へ寄与する2つの側面を示唆している。1つ目は、蓋然性の高い視聴覚刺激の対応づけを選択することであり、2つ目は、蓋然性の低い視聴覚刺激の対応づけを棄却するというものである。複数の対応の候補から素早く1つの答えを導くために、合わない要素を持つ対応づけの候補を積極的に棄却しているのかもしれない。

## 引用文献

- Bertelson, & Aschersleben. (1998). *Psychonomic Bulletin & Review*, 5, 482-489.
- Kanaya, & Yokosawa. (2011). *Psychonomic Bulletin & Review*, 18, 123-128.
- Schutz, & Kubovy. (2009). *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1791-1810.
- Vatakis, & Spence. (2008). *Acta Psychologica*, 127, 12-23.
- Vroomen, & de Gelder. (2004). *Handbook of Multisensory Processes* (pp. 141-150).